View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Vision Research 51 (2011) 2145–2155

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Exploring the time course of face matching: Temporal constraints impair unfamiliar face identification under temporally unconstrained viewing

Müge Özbek [a,c], Markus Bindemann [b,c,*]

[a] *Department of Psychology, Boğaziçi University, Turkey*
[b] *School of Psychology, University of Kent, CT2 7NP, UK*
[c] *Department of Psychology, University of Essex, UK*

ARTICLE INFO

Article history:
Received 6 May 2011
Received in revised form 27 July 2011
Available online 16 August 2011

Keywords:
Face perception
Unfamiliar faces
Time course
Eye movements

ABSTRACT

The identification of unfamiliar faces has been studied extensively with matching tasks, in which observers decide if pairs of photographs depict the same person (identity matches) or different people (mismatches). In experimental studies in this field, performance is usually self-paced under the assumption that this will encourage best-possible accuracy. Here, we examined the temporal characteristics of this task by limiting display times and tracking observers' eye movements. Observers were required to make match/mismatch decisions to pairs of faces shown for 200, 500, 1000, or 2000 ms, or for an unlimited duration. Peak accuracy was reached within 2000 ms and two fixations to each face. However, intermixing exposure conditions produced a context effect that generally reduced accuracy on identity mismatch trials, even when unlimited viewing of faces was possible. These findings indicate that less than 2 s are required for face matching when exposure times are variable, but temporal constraints should be avoided altogether if accuracy is truly paramount. The implications of these findings are discussed.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

In unfamiliar face matching, observers have to decide if two simultaneous presentations of a face belong to the same person or different people. This task can be remarkably difficult. Investigations into the utility of photo-identity cards, for example, have revealed more than 50% errors in face matching under challenging task demands (Bindemann & Sandford, 2011; Kemp, Towell, & Pike, 1997), and observers continue to average 10–30% errors under optimized laboratory conditions (see, e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010; Megreya, Bindemann, & Havard, 2011). In everyday functioning, observers may be unaware of such failures (e.g., Scheck, Neufeld, & Dwyer, 2000; Simons & Levin, 1998) but misidentifications do occur and can have serious consequences at an individual level and beyond (Davies & Griffiths, 2008; Wells, Memon, & Penrod, 2006). The study of face matching is therefore a topic of theoretical and practical importance.

Many of the factors that give rise to failures in face matching are now well understood. Every encounter with a person provides a different pattern for visual analysis, due to, for example, changes in lighting, viewpoint, facial expression, and, in the case of photographs, distortion by a camera lens (e.g., Bruce, 1982; Johnston, Hill, & Carman, 1992; Logie, Baddeley, & Woodhead, 1987; Longmore, Liu, & Young, 2008). These contextual variables influence the appearance of a face and, as a result, the identification of unfamiliar faces suffers. Moreover, this difficulty increases continuously the more viewing conditions are compromised (Longmore, Liu, & Young, 2008). Unfamiliar face identification is therefore a taxing process that is highly susceptible to such task demands.

To explore these difficulties in face matching, performance is usually self-paced in psychological experiments under the assumption that this will encourage best-possible accuracy (see, e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010; Megreya, Bindemann, & Havard, 2011). As a consequence, however, the temporal characteristics of this task remain largely unexplored. This is remarkable for several reasons. From a practical perspective, face matching in operational contexts, such as passport control, is a routine task that can be subject to time limits depending on demand. It is therefore valuable to establish the minimum time required to perform this task reliably. From a theoretical perspective, on the other hand, the temporal characteristics of visual tasks are linked intimately to the underlying cognitive processes (e.g., Henderson, 2003, 2007). The time course of face matching may therefore provide important insights regarding the processes by which this task is solved.

While current knowledge of the time course of face matching is limited, studies of *recognition memory* provide some clues

* Corresponding author at: School of Psychology, University of Kent, CT2 7NP, UK. Fax: +44 (0)1227 827030.
E-mail address: m.bindemann@kent.ac.uk (M. Bindemann).

0042-6989/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2011.08.009

regarding the temporal characteristics of this task. It has been found, for example, that newly learned faces can be recognized surprisingly quickly when viewing is restricted. In these studies, viewing conditions were limited temporally by restricting display times to predefined exposure durations (Meinhardt-Injac, Persike, & Meinhardt, 2010; Veres-Injac & Schwaninger, 2009), or by monitoring eye movements to limit the acquisition of facial information to a set number of fixations (Hsiao & Cottrell, 2008). These studies show that maximal identification accuracy can be reached with exposure durations of only 90 ms to a face (Veres-Injac & Schwaninger, 2009) or with as little as two fixations (Hsiao & Cottrell, 2008). These findings therefore provide an interesting contrast to the difficulties that have been documented widely in face matching by indicating that observers, despite the seemingly challenging demands of this task, might not benefit from viewing unfamiliar faces for even slightly extended durations.

However, these studies differ from matching paradigms in two important ways. In these experiments, observers were required to recognize a face after an opening study phase of 1.5 and 3 s exposure to a face (Hsiao & Cottrell, 2008; Veres-Injac & Schwaninger, 2009). Rapid identification at test therefore occured following more extensive exposures during initial face encoding. In a sense, these paradigms are therefore comparable to studies of familiar face processing, which show that faces can be recognized quickly once they have been learned (e.g., Ellis, Young, & Koenken, 1993; Morrison, Bruce, & Burton, 2000; Young, Hellawell, & de Haan, 1988), rather than the identification of unfamiliar faces *per se*. Moreover, because recognition memory studies do not require the simultaneous presentation of two face images, identification was probed in these experiments across the same face image. This is advantageous under some conditions, for example, for exploring the contribution of specific stimulus features to face perception (see Meinhardt-Injac, Persike, & Meinhardt, 2010; Veres-Injac & Schwaninger, 2009). However, under these conditions it also becomes possible to identify a person by encoding specific pictorial properties of a face image, in a manner in which the same image of *any* visual object could be remembered, rather than by processing mechanisms that require specific face perception skills (see, e.g., Bruce, 1982; Duchaine & Nakayama, 2004; Duchaine & Weidenfeld, 2003; Longmore, Liu, & Young, 2008). It therefore remains unresolved whether unfamiliar face processing exhibits a similarly short time course in matching tasks, in which different images of the same faces are presented simultaneously for person identification.

To investigate these issues, this study presented observers with pairs of faces comprising photographs of the same person or two different people, and match/mismatch decisions to these facial identities were required (as in, e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010; Megreya, Bindemann, & Havard, 2011). In this task, different photographs of the same person were provided on identity-match trials to eliminate simple picture-matching strategies (see Bruce, 1982; Longmore, Liu, & Young, 2008) and both images in a pair were always shown simultaneously to avoid the prior familiarization with a face. To examine the time course of face matching, exposure to these face pairs was then controlled by limiting display times to four different exposure durations, set at 200, 500, 1000, and 2000 ms. A fifth condition allowed unlimited viewing of the face stimuli, to provide a baseline for best-possible accuracy. With these manipulations, we sought to investigate if peak identification accuracy requires self-paced face viewing in line with the demands of previous matching tasks (e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010; Megreya, Bindemann, & Havard, 2011) or conversely, if this is already possible with limited display durations.

To study how any identification decisions are achieved in more detail, eye movements were also recorded in this task. Observers' eye movements provide a real-time index of ongoing cognitive processing (e.g., Henderson, 2003, 2007) and have been used widely to study face perception (see, e.g., Althoff & Cohen, 1999; Bindemann, Scheepers, & Burton, 2009; Blais et al., 2008; Haith, Bergman, & Moore, 1977; Janik et al., 1978). In unfamiliar face processing, eye movements are clearly important for visual encoding, as person recognition is impaired when scanning behavior is restricted during initial face learning (Henderson, Williams, & Falk, 2005). At the same time, as few as two fixations to a face can suffice for recognition after a face has been learned (Hsiao & Cottrell, 2008). This demonstrates that eye movements are important for the encoding of unfamiliar faces, but also suggests that only a few looks are required for identification under some conditions. In this study, we therefore tracked observers' eye movements to determine how many fixations are required for accurate person identification in a face-matching task and whether specific viewing strategies are employed. For this purpose, 200 ms was selected as the shortest display time, as this is too brief to permit fixation of the face stimuli (Hallett, 1986). Longer display durations then provide increasing opportunities to fixate the face pairs and to encode these stimuli in detail.

2. Experiment 1

2.1. Method

2.1.1. Participants

Thirty students (14 males) from the University of Essex, with a mean age of 25.8 years (SD = 5.9), volunteered for this experiment for a small fee. All had normal vision.

2.1.2. Stimuli

The stimuli consisted of 200 face pairs from the Glasgow University Face Database (GUFD; see Burton, White, & McNeill, 2010). Half of the pairs depicted identity *matches*, in which two different photographs of the same person were shown, while the remaining face pairs were identity *mismatches*, in which two different people were depicted. In addition, these face pairs were split evenly to depict male and female face stimuli. The pairs were constructed so that all faces were shown in grayscale on a white background. Each face measured maximally 260 pixels in width at a screen resolution of 66 pixels/in. At a viewing distance of 80 cm, held constant by a chin rest, this equates to ~7.2° of visual angle (VA). The faces in each pair were arranged in such a way that the horizontal distance between the center of each face measured 372 pixels (~10.2° VA).

In each *match* and *mismatch* pair, one face image was taken with a high-quality digital camera while the other was a frame of a person's face from high-quality video footage. For each person in the GUFD, these images were taken only a few minutes apart, in the same full-face pose, with a neutral expression, and under the same lighting conditions. On identity *match* trials, these photographs therefore provided similar but not identical images of a person to ensure that the task cannot be done using simple image matching processes (for further details about the stimuli, see Burton, White, & McNeill, 2010). For identity *mismatches*, on the other hand, the GUFD provides perceived-similarity ratings for all faces in the database, which were used to construct *mismatch* pairs with the highest possible similarity. The average similarity rating for the 100 *mismatch* pairs was 0.42/1 (SD = 0.06). For example stimuli, see Figs. 5 and 6.

2.1.3. Procedure

The stimuli were displayed using SR-Research Experiment-Builder software (Version 1.4.2) on a 21 in. color monitor that was connected to an SR-Research Eyelink 1000 desk-mounted eye tracking system running at 500 Hz sampling rate. Viewing was binocular, but only the participants' left eye was tracked. To calibrate the tracker, participants fixated a series of nine fixation targets on the display monitor. Calibration was then validated against a second sequence of nine fixation targets, and if the latter indicated poor measurement accuracy, calibration was repeated. This procedure was carried out at the beginning of the experiment and every 40 trials thereafter.

Each trial began with the presentation of a central dot, which participants were asked to fixate so that an automatic drift correction could be performed. Once participants fixated this dot, the experimenter pressed a button to initiate a trial. A face pair was then presented for 200, 500, 1000 or 2000 ms and was then replaced by a blank screen, or remained onscreen until a response was registered, in the unlimited viewing condition. Participants were informed of the purpose of the experiment in advance, including the different exposure conditions. As such, they were instructed to decide whether an identity *match* or *mismatch* was shown in each trial, by using their left and right index fingers to press the corresponding keys on a button box. Accuracy was emphasized and responses were self-paced. Thus, participants were informed that they could respond while the faces were still onscreen or, for the short display durations, even after the face had disappeared from view.

Each participant was given 200 experimental trials, consisting of 20 *match* and 20 *mismatch* trials for each of the display conditions (200, 500, 1000, 2000 ms, and unlimited). The stimulus set was rotated around conditions so that each face pair was only shown once to each participant in any of the conditions. However, over the course of the experiment the presentation of face pairs was counterbalanced across participants, so that each stimulus appeared in each condition an equal number of times. The presentation of the conditions was randomly intermixed throughout the task and participants were given short breaks every 40 trials, followed by a re-calibration phase.

2.2. Results

2.2.1. Accuracy

The data of one participant who performed near chance (54% accuracy) throughout the experiment was removed from the analysis. Across the remaining participants, accuracy for *match* and *mismatch* decisions was then analyzed for the five exposure durations. The mean percentage accuracy for these conditions is illustrated in Fig. 1. The combined data for *match* and *mismatch* trials shows that accuracy increased with exposure duration and peaked in the 2000 ms condition. In line with these observations, a one-factor (200, 500, 1000, 2000 ms, unlimited) ANOVA of this data showed a main effect of exposure duration, $F(4,112) = 57.73$, $p < 0.0001$. Tukey HSD test revealed that accuracy increased with each exposure condition, all $qs \geqslant 3.97$, all $ps < 0.05$, except from the 2000 ms to the unlimited displays, $q = 0.00$, $p > 0.05$.

In a next step, accuracy was broken down into *match* and *mismatch* trials. This data shows that performance on *match* trials was close to the chance level of 50% in the 200 ms condition, but then improved substantially with a 500 ms exposure. *Match* performance increased further with even longer exposures, but appeared to be similar in the 1000 ms, 2000 ms and the unlimited viewing condition. In contrast, *mismatch* accuracy was substantially higher than *match* accuracy in the 200 ms condition, indicating a possible *mismatch* response bias at these short exposure times. However, *mismatch* accuracy then showed a less marked improvement with longer exposure durations and, with the exception of the 200 ms condition, was generally lower than on *match* trials.

To analyze the accuracy data for *match* and *mismatch* trials, a 2×5 within-subjects ANOVA of trial type (*match*, *mismatch*) and exposure duration (200, 500, 1000, 2000 ms, unlimited) was conducted. ANOVA did not show a main effect of trial type, $F(1,28) = 2.51$, $p = 0.12$, but revealed a main effect of exposure duration, $F(4,112) = 57.73$, $p < 0.0001$, and an interaction between both factors, $F(4,112) = 16.94$, $p < 0.0001$. This interaction was qualified by simple main effects of exposure duration for *match*, $F(4,112) = 84.88$, $p < 0.0001$, and *mismatch* pairs, $F(4,112) = 4.93$, $p < 0.01$. For *match* trials, Tukey HSD test showed that accuracy was lower in the 200 ms condition than for any other exposure time, all $qs \geqslant 14.55$, all $ps < 0.001$. Similarly, accuracy for the 500 ms condition was worse than for 1000 ms, 2000 ms, and unlimited displays, all $qs \geqslant 5.23$, all $ps < 0.01$. In addition, *match* performance was similar for 1000 ms, 2000 ms and unlimited displays, all $qs \leqslant 2.19$, all $ps > 0.05$. A different pattern emerged on *mismatch* trials. Here, accuracy for the 200, 500 and 1000 ms conditions was indistinguishable, all $qs \leqslant 0.35$, all $ps > 0.05$, but observers were slightly more accurate in the 2000 ms and unlimited conditions. These differences were reliable between the 200, 500, 1000 ms displays compared to the 2000 ms condition, all $qs \geqslant 3.99$, all $ps < 0.05$. Similarly, the unlimited displays yielded more accurate responses than 500 ms exposures, $q = 4.09$, $p < 0.05$. The differences between the 200 ms and the unlimited condition, $q = 3.90$, $p > 0.05$, and between the 1000 ms and the

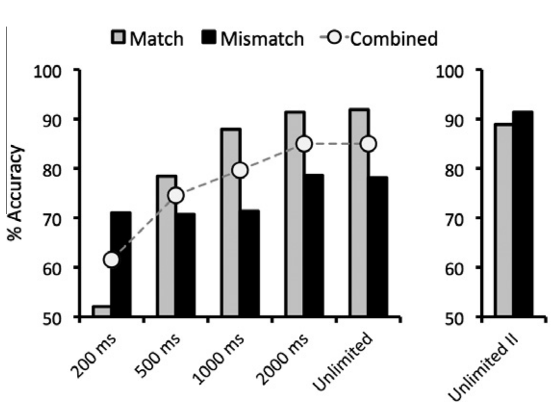

Fig. 1. Mean accuracy for the match and mismatch conditions in Experiment 1 (for 200, 500, 1000, 2000 ms, and unlimited displays) and in Experiment 2 (unlimited II). Circles indicate overall matching performance (the average of match and mismatch accuracy).

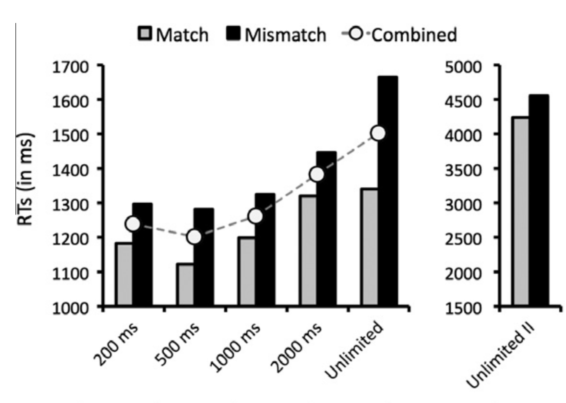

Fig. 2. Mean reactions times for the match and mismatch conditions in Experiment 1 (for 200, 500, 1000, 2000 ms, and unlimited displays) and in Experiment 2 (unlimited II). Circles indicate overall matching performance (the average of match and mismatch accuracy).

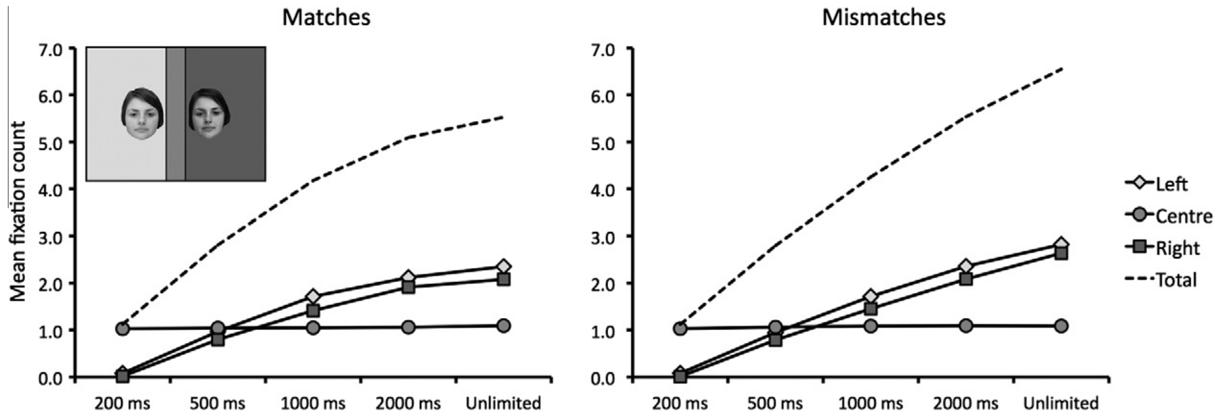

Fig. 3. Mean number of fixations to the central ROI, and to the left and right face in a stimulus pair for the experimental conditions. Total represents the sum of fixations to all ROIs. The color-coded inset illustrates the area of the three ROIs.

unlimited condition, $q = 3.71$, $p > 0.05$, were also approaching significance.

Taken together, these accuracy data point to an early bias to classify faces as *mismatches* in the 200 ms condition, which leads to near chance performance on *match* trials. Longer intervals, however, are marked by an advantage for *match* trials, which might reflect a bias to respond *same* to the face pairs. These observations are confirmed when this data is couched as detection measures, which gives d' values of 1.32, 2.23, 2.66, 3.06, 3.09 and a criterion of 0.31, −0.12, −0.32, −0.28, and −0.29 for the 200, 500, 1000, 2000 ms and unlimited condition, respectively. A one-factor ANOVA of the d' data reveals a main effect of condition, $F(4,112) = 57.76$, $p < 0.0001$, reflecting differences between all conditions, all $qs \geqslant 4.24$, all $ps < 0.05$, except the 2000 ms and unlimited displays, $q = 0.27$, $p > 0.05$. Analogous analysis of criterion, $F(4,112) = 19.22$, $p < 0.0001$, shows that observers applied a different classification criterion in the 200 ms condition in comparison with all other conditions, all $qs \geqslant 7.05$, all $ps < 0.001$, while the four longer exposure conditions did not differ from each other, all $qs \leqslant 3.41$, all $ps > 0.05$. These measures therefore indicate that the ability to discriminate *match* and *mismatch* pairs improves with increasing exposure time to peak with 2000 ms displays. However, observers exhibit an initial *mismatch* response bias, with 200 ms displays, which is then replaced by a *match* bias in all other conditions.

2.2.2. Reaction times

Although task instructions emphasized accuracy, the mean correct response times were also analyzed for *match* and *mismatch* trials (see Fig. 2). As with accuracy, a one-factor ANOVA of the combined reaction times (*match* and *mismatch*) revealed an effect of exposure duration, $F(4,112) = 10.61$, $p < 0.0001$. Tukey HSD test showed that response times were similar for the 200, 500 and 1000 ms conditions, all $qs \leqslant 1.58$, all $ps > 0.05$, and were faster in these conditions compared to unlimited displays, all $qs \geqslant 6.34$, all $ps < 0.001$. In addition, responses were also reliably slower in the 2000 ms condition compared to 500 ms displays, $q = 4.78$, $p < 0.01$. None of the remaining comparisons were significant, all $qs \leqslant 3.79$, all $ps > 0.05$.

As with the accuracy data, response times were also broken down into *match* and *mismatch* conditions. A 2 (*match*, *mismatch*) × 5 (200, 500, 1000, 2000 ms, unlimited) within-subjects ANOVA of this data revealed a main effect of trial type, $F(1,28) = 11.07$, $p < 0.01$, due to longer response times on *mismatch* trials, and a main effect of exposure time, $F(4,112) = 10.61$, $p < 0.001$. The follow-up comparisons for the main effect of exposure are identical to the combined analysis of the *match* and *mismatch* data (see above). The interaction between both factors failed to reach significance, $F(4,112) = 2.27$, $p = 0.07$.

Thus, the reaction time data shows that observers required more time for *mismatch* than *match* decisions. Unsurprisingly, participants also took more time to respond with unlimited displays. It is notable, however, that even in the 2000 ms and the unlimited conditions, responses were made on average within 1.7 s of stimulus onset. This finding converges with the accuracy data to suggest that maximum performance is reached within 2000 ms.

2.2.3. Eye movements

Observers' eye movements during face matching were examined for all correct trials.[1] Eye movements were preprocessed by integrating fixations of less than 80 ms with the immediately preceding or following fixation if that fixation lay within half a degree of visual angle. Otherwise these short fixations were excluded. The rationale for this was that such short fixations usually result from false saccade planning and are unlikely to reflect meaningful information processing (see Rayner & Pollatsek, 1989). The fixation data were then compared to a set of three predefined regions of interest (ROIs), reflecting the central area between both faces, and the location of the left and the right face in each face pair. These were defined broadly into a central vertical region measuring 108 (W) × 768 (H) pixels and equivalent to 10.6% of the total area of the displays, and two flanking regions measuring 458 (W) × 768 (H) pixels and equivalent to 44.7% of the display.

Fig. 3 illustrates the average number of fixations that were made to these ROIs during the presentation of the *match* and *mismatch* face pairs in the five exposure conditions. This data shows that only a single fixation, which was located in the central region of the screen, was recorded on *match* and *mismatch* trials of the 200 ms condition. Note that the stimulus displays were preceded by a central fixation point and this initial fixation simply reflects the location of observers' eyes at stimulus onset. It is therefore referred to as Fixation 0 in the analysis, as it is not driven by a saccade to regions of interest within the stimulus displays. A preliminary analysis showed that this central region received an equal number of fixations across all conditions (mean = 1.06 fixations, SD = 0.03). These initial fixations on the central ROI were therefore omitted from all subsequent analyses.

Longer exposure durations then enabled observers to fixate the face stimuli. Fig. 3 shows that the 500 ms condition permitted observers to fixate each face in a stimulus pair once, and the num-

[1] Eye movements on incorrect trials yielded similar results across all measures reported here, but are not reported for brevity. This data is available on request from the authors.

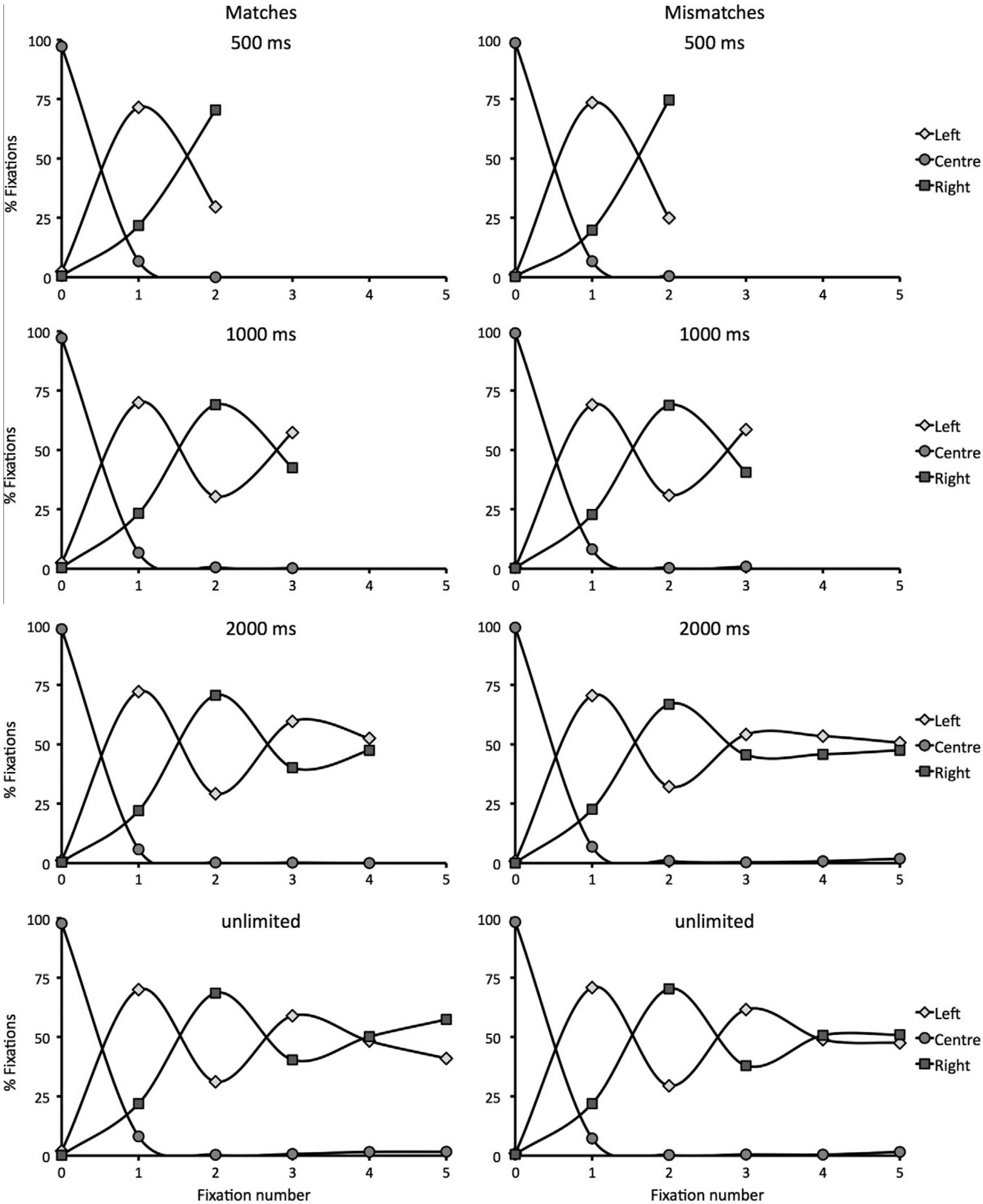

Fig. 4. Mean percentage fixations to the left, central, and right ROI for the 500, 1000, 2000 ms and the unlimited condition. Up to the first five fixations to the stimulus displays are shown.

ber of fixations to the faces then continued to increase with exposure duration. A 2 (*match*, *mismatch*) × 2 (left, right ROI) × 4 (500, 1000, 2000 ms, unlimited) within-subjects ANOVA of this data revealed a main effect of ROI, $F(1,28) = 21.88$, $p < 0.0001$, reflecting more fixations to the left than to the right ROI. In addition, a main effect of trial type, $F(1,28) = 11.04$, $p < 0.01$, of duration, $F(3,84) = 81.34$, $p < 0.0001$, and an interaction between both factors was found, $F(3,84) = 8.49$, $p < 0.0001$. Simple main effect analysis showed an effect of trial type for the unlimited condition, $F(1,28) = 22.32$, $p < 0.0001$, but not for the 500, 1000 and 2000 ms displays, all $Fs \leqslant 1.22$. In addition, simple main effects of duration were found for *match*, $F(3,84) = 29.29$, $p < 0.0001$, and *mismatch* trials, $F(3,84) = 54.51$, $p < 0.0001$. For both types of face pairs, Tukey HSD test revealed reliable differences between all the exposure conditions, all $qs \geqslant 4.14$, all $ps < 0.05$, except between 2000 ms and unlimited *match* displays, $q = 1.84$, $p > 0.05$.

Taken together, these eye movements reveal that only a single fixation was made with a 200 ms display time, which was located in the central region of the display. With longer exposures, observers were able to fixate the faces directly. In the 500 ms condition, each face received approximately one fixation and, unsurprisingly, this number increased further with exposure duration.

2.2.4. Viewing strategy

The percentages of each fixation that were directed at the left, central and right ROI were also broken down by individual fixations (i.e., the 1st, 2nd, 3rd, etc. fixation that was made to each display) to visualize observers' viewing strategies in the matching task (see Fig. 4). For all conditions, this data shows a strong tendency to initially fixate the left face in a display, with the 1st fixation, followed by an immediate switch to the right face on the 2nd fixation. With 1000, 2000 ms and unlimited displays, this is followed by a further switch back to the left face. In combination with the accuracy data, this suggests that two or three fixations to the faces are sufficient to reach optimal performance on *match* trials or, effectively, one fixation to each face plus one additional fixation. For *mismatch* trials, on the other hand, observers required more time to reach peak accuracy and, on average, two to three fixations to *each* face.

2.2.5. Distribution of fixations across faces

Finally, the distribution of fixations was also plotted to illustrate where eye movements were directed in a face. For this purpose, all fixations were fitted with a Gaussian and the *z*-scored distribution of these Gaussians was plotted for the experimental conditions (for similar analysis, see Bindemann, 2010; Blais et al., 2008). For *match* trials these distributions are illustrated in Fig. 5 and for *mismatch* trials in Fig. 6. These fixation distributions converge with the viewing strategies plotted in Fig. 4 by showing that observers tended to initially view the left face in a stimulus pair, then switched to the right face, before switching back to the face on the left. In addition, these distributions show that fixations generally clustered around the central region of a face, encompassing the eye regions and the nose.

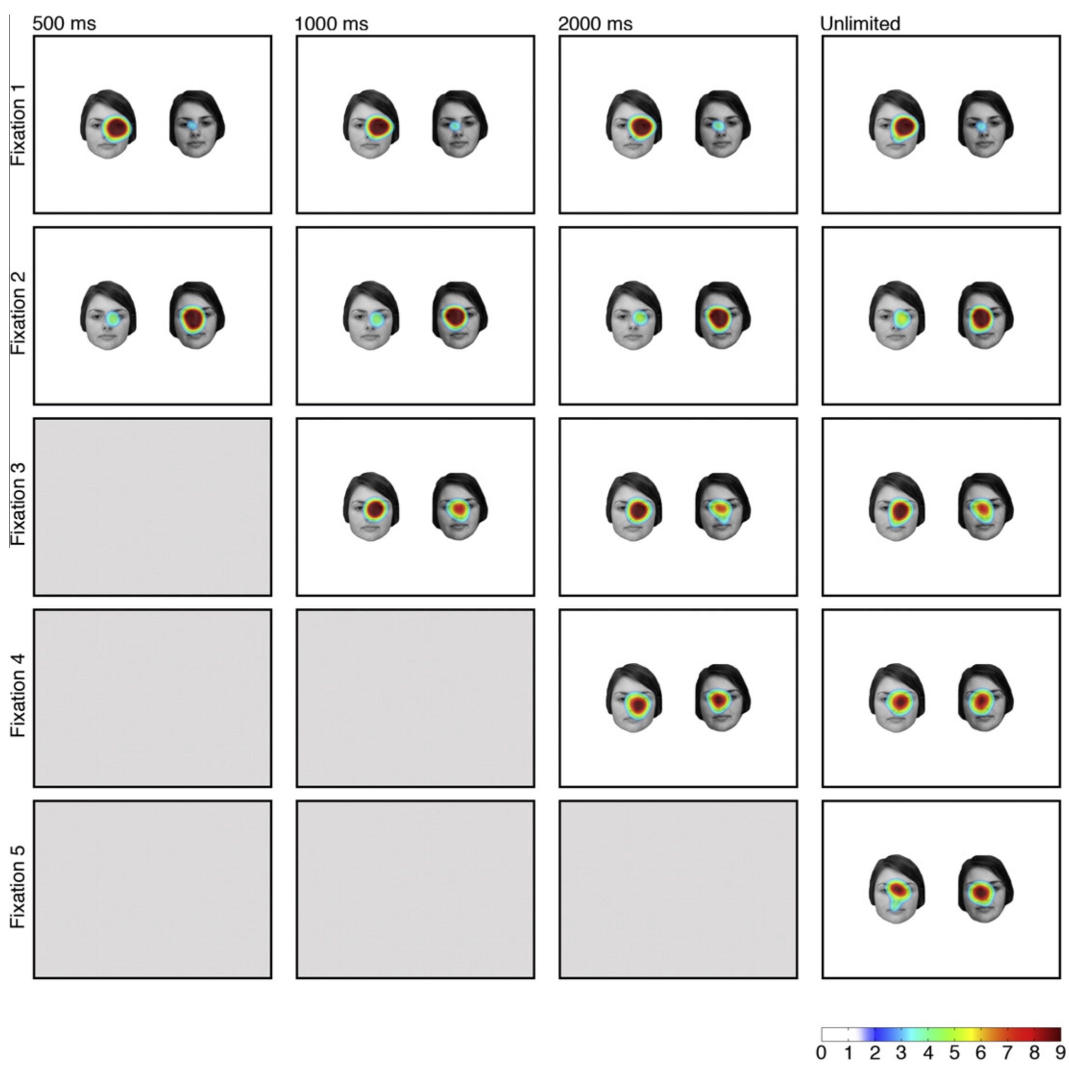

Fig. 5. Gaussian maps illustrating the *z*-scored distribution of fixations in match displays for the 500, 1000, 2000 ms and the unlimited condition, superimposed on an example of a match stimulus. (For interpretation to colours in this figure, the reader is referred to the web version of this paper.)

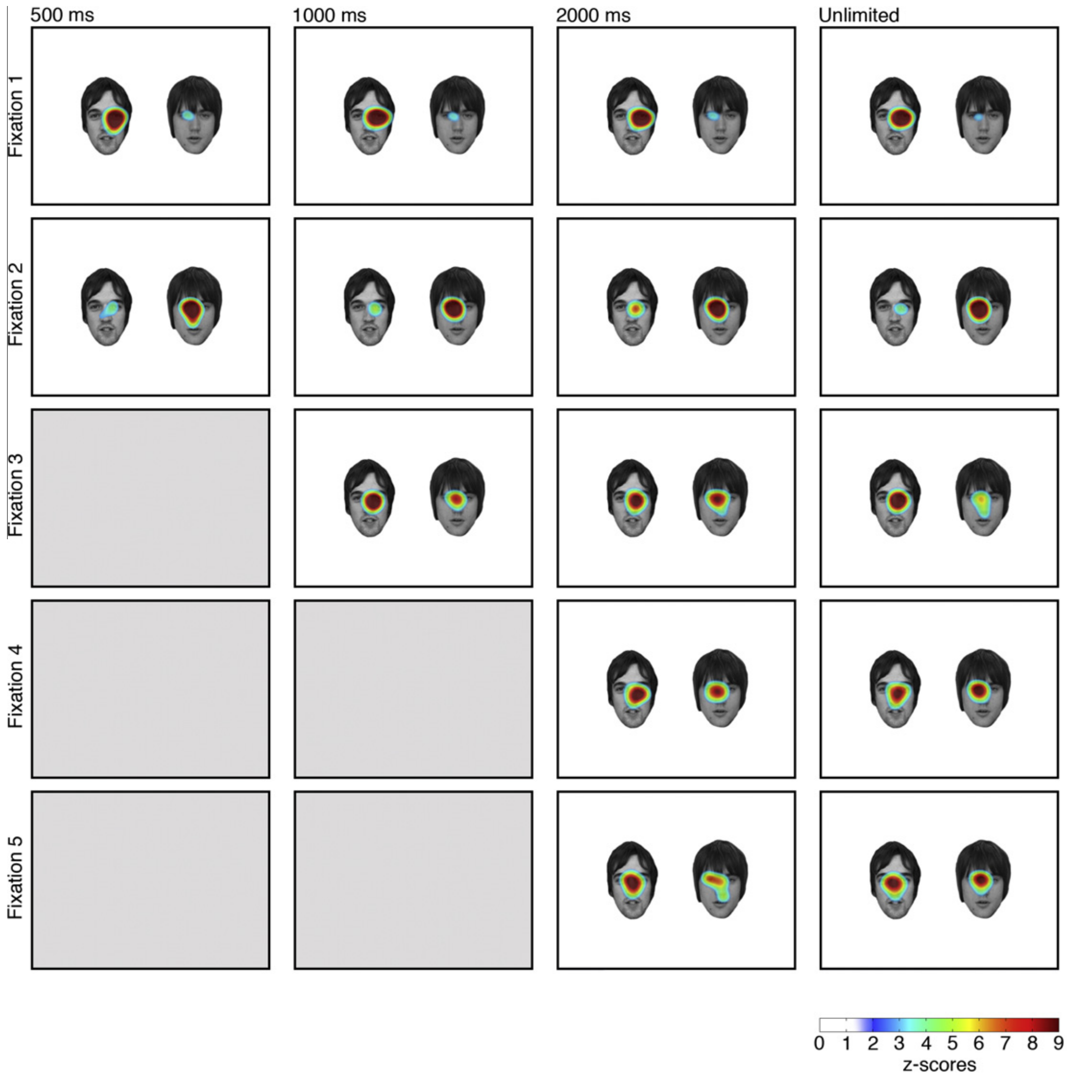

Fig. 6. Gaussian maps illustrating the *z*-scored distribution of fixations in mismatch displays for the 500, 1000, 2000 ms and the unlimited condition, superimposed on an example of a mismatch stimulus. (For interpretation to colours in this figure, the reader is referred to the web version of this paper.)

2.3. Discussion

This experiment examined the time course of face-matching by constraining viewing conditions temporally and by tracking observers' eye movements. The results show that *match* identification was close to chance with a 200 ms exposure, while *mismatch* accuracy was comparatively higher in this condition, at over 70% correct. However, while *match* accuracy increased rapidly thereafter, and reached maximum accuracy with a 1000 ms viewing time, *mismatch* accuracy increased more gradually and peaked later, in the 2000 ms condition. These findings are also reflected in observers' eye movements, which reveal an efficient viewing strategy. This strategy involved sampling the left face in a stimulus pair with the 1st fixation, followed by a single fixation to the right face, before observers returned to the left face with the 3rd saccade. On *match* trials, these three fixations appeared sufficient to achieve optimal accuracy. On *mismatch* trials, on the other hand, the same viewing strategy was employed but further fixations were needed to make a decision.

In addition to these general observations, several interesting findings emerge from this experiment. Firstly, the accuracy data points to an early bias to classify face pairs as identity *mismatches* in the 200 ms condition. In this condition, this led to near-chance performance on *match* trials, while accuracy was reliably higher in the *mismatch* condition. It seems likely that this bias arises from the use of different face photographs of each facial identity for *match* trials and because the 200 ms condition provides insufficient time to fixate the face stimuli (Hallett, 1986). As a consequence, these images may inherently look dissimilar, and are therefore perceived as identity *mismatches*, when viewing is compromised in this manner. The remarkable aspect of this early *mismatch* bias is, however, that *mismatch* accuracy improved much less at longer intervals compared to *match* trials. As a result, the remaining conditions yielded an advantage to detect *match* pairs in this experimental setting.

These response biases differ markedly from other studies in this field, which have consistently shown equivalent performance for *match* and *mismatch* trials (see, e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010; Megreya, Bindemann, & Havard, 2011). However, previous studies only assessed face-matching performance under temporally unconstrained viewing conditions. This suggests that the discrepancy with previous findings arises from the current experimental context, in which different exposure durations were intermingled. This context appears to produce response biases in the temporally constrained conditions that spread to the unlimited displays. This notion receives some further support from the reaction times, which show that observers failed to take full advantage of the 2000 ms and unlimited displays by responding in less than the available viewing time in these experimental conditions.

To investigate this issue directly, a control experiment was conducted in which face matching was assessed under unconstrained viewing only. The aim of this experiment was to establish whether *mismatch* performance is more evenly matched with *match* accuracy under these specific conditions, in line with previous studies in this field (e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010). This experiment is therefore essentially a simple replication of these previous studies, to confirm that the response biases in Experiment 1 reflect the combination of exposure conditions, rather than perhaps a spurious aspect of our stimuli and procedure.

3. Experiment 2

3.1. Method

3.1.1. Participants

Twenty new students (8 males) from the University of Essex, with a mean age of 25.0 years (SD = 4.4), took part in this experiment for a small fee. All had normal vision.

3.1.2. Stimuli and procedure

Stimuli and procedure were identical to Experiment 1, except that all face stimuli were presented under temporally unconstrained viewing conditions. Each participant therefore received 100 *match* and 100 *mismatch* trials. The presentation of these stimuli was randomly intermixed throughout and participants were given short breaks every 40 trials. As in Experiment 1, participants were informed of the purpose of the experiment in advance and accuracy was emphasized.

3.2. Results and discussion

3.2.1. Accuracy and reaction times

As in Experiment 1, accuracy and mean reaction times were calculated for *match* and *mismatch* trials and are illustrated in Figs. 1 and 2. A paired-samples *t*-test of this data shows that accuracy was comparable for *match* and *mismatch* trials, $t(19) = 0.73$, $p = 0.48$, and reaction times were also similar for both conditions, $t(19) = 0.72$ $p = 0.48$. To compare this data with the unlimited viewing condition of Experiment 1, a 2 (Exp. 1 vs. Exp. 2) × 2 (*match*, *mismatch*) mixed-factor ANOVA was conducted for the accuracy data, which showed an interaction between both factors, $F(1,47) = 10.08$, $p < 0.01$. Simple main effect analysis showed an effect of trial type for Experiment 1, $F(1,47) = 17.69$, $p < 0.001$, but not for Experiment 2, $F(1,47) < 1$, and also an effect of experiment for *mismatch* trials, $F(1,94) = 14.90$, $p < 0.001$, but not for *match* trials, $F(1,94) < 1$. In addition, an analogous ANOVA of reaction times revealed a main effect of experiment, $F(1,47) = 238.49$, $p < 0.0001$, reflecting longer reaction times in Experiment 2, but no main effect of condition was found, $F(1,47) = 2.88$, $p = 0.10$, and no interaction between both factors, $F(1,47) < 1$. Finally, we also compared the detection measures d' and *criterion* for the accuracy data of Experiments 1 and 2. This analysis showed that observers were more accurate in the unlimited condition of Experiment 2 ($d' = 3.09$ and 3.51 for Experiments 1 and 2, respectively, $t(47) = 2.21$, $p < 0.05$), and no longer exhibited a response bias to classify the face pairs as *same* in this study (*criterion* = −0.29 and 0.06 for Experiments 1 and 2, $t(47) = 3.05$, $p < 0.01$).

3.2.2. Eye movements

The differences seen in the behavioral data were also reflected in observers' eye movements for the correct trials. On average, observers made 14.9 fixations on *match* trials and 15.4 fixations on *mismatch* trials in Experiment 2 (not including Fixation 0). To compare this data with the unlimited condition of Experiment 1, a 2 (Exp. 1 vs. Exp. 2) × 2 (*match*, *mismatch*) mixed-factor ANOVA was conducted, which found no main effect of trial type, $F(1,47) = 1.41$, $p = 0.24$, and no interaction between both factors, $F(1,47) < 1$. However, a main effect of experiment was found, $F(1,47) = 21.30$, $p < 0.001$, which shows that substantially more fixations were made in Experiment 2.

These fixations were evenly distributed between the left faces (average number of fixations on *match* trials = 7.6, *mismatch* trials = 7.7) and the right faces (*match* trials = 7.3, *mismatch* trials = 7.6) of the stimulus pairs. In line with these observations, a 2 (*match*, *mismatch*) × 2 (left, right) ANOVA showed no main effect of trial type, $F(1,19) < 1$, or ROI, $F(1,19) = 2.34$, $p = 0.14$, and no interaction between both factors, $F(1,19) = 3.29$, $p = 0.09$. A more detailed illustration of this data reveals a viewing strategy that is similar to Experiment 1 (see Fig. 7). Thereby, initial fixations were predominantly directed at the left face in a stimulus pair, then shifted to the face on the right, before the number of fixations that each face received gradually converged around the 50% mark. In contrast to Experiment 1, however, these viewing strategies operate along a different time course, whereby the initial switching between the left and the right face occurred more gradually (cf. Figs. 4 and 7).

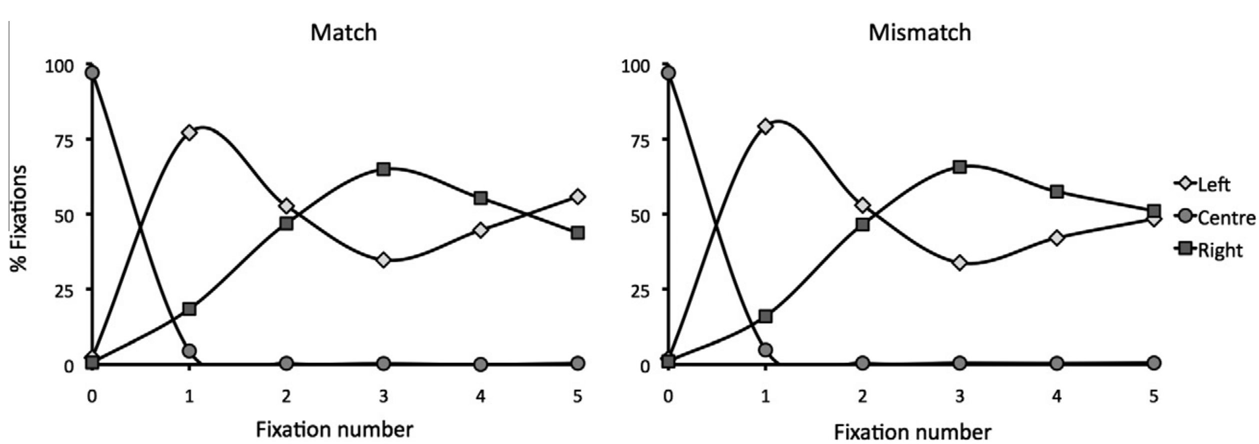

Fig. 7. Mean percentage fixations to the left, central, and right ROI in Experiment 2. The first five fixations to the stimulus displays are shown.

This impression is reinforced by an illustration of the distribution of fixations (see Fig. 8) and difference maps for the unlimited conditions, which were calculated by subtracting the distribution of fixations in Experiment 2 from Experiment 1 (for similar analyses, see Bindemann, 2010; Blais et al., 2008). These difference maps show clearly that both experiments elicited different viewing strategies. Moreover, this data also suggests that observers sampled more facial features in Experiment 2. For example, while observers predominantly focused on the inner eyes of the face pairs in Experiment 1 (i.e., the right eye of the left face, and the left eye of the right face), the outer eyes were fixated more directly in Experiment 2. Similarly, while observers largely neglected the mouth regions in Experiment 1, these regions received more fixations in Experiment 2.

Taken together, this data shows that observers were equally accurate on *match* and *mismatch* trials in Experiment 2, and more accurate than in Experiment 1, but also took longer to categorize the face pairs in both conditions. In addition, observers looked at the faces more often in Experiment 2 and their eye movements were characterized by a more gradual switch between both faces in a pair, indicating that these stimuli were encoded more thoroughly than in Experiment 1. This control study therefore provides confirmatory evidence that intermixing different exposure conditions induced a context effect in Experiment 1, whereby the temporally constrained viewing of the face pairs also speeded identification decisions under unlimited viewing and reduced *mismatch* accuracy.

4. General discussion

Matching unfamiliar faces according to identity is an error-prone task (e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010). To maximize performance, these tasks are therefore usually conducted under self-paced conditions that emphasize accuracy. The current study contravened this convention to explore the time course of face matching. Specifically, we sought to explore if maximum accuracy requires self-paced viewing of face stimuli or if peak performance can be achieved when viewing is restricted. Furthermore, we tracked observers' eye movements to determine whether specific viewing strategies are employed during face matching.

With only a 200 ms exposure to pairs of faces, identification was near chance on identity *match* trials, while accuracy was improved for identity *mismatches*. However, we attribute this *mismatch*-to-*match* advantage to a response bias that arises from the

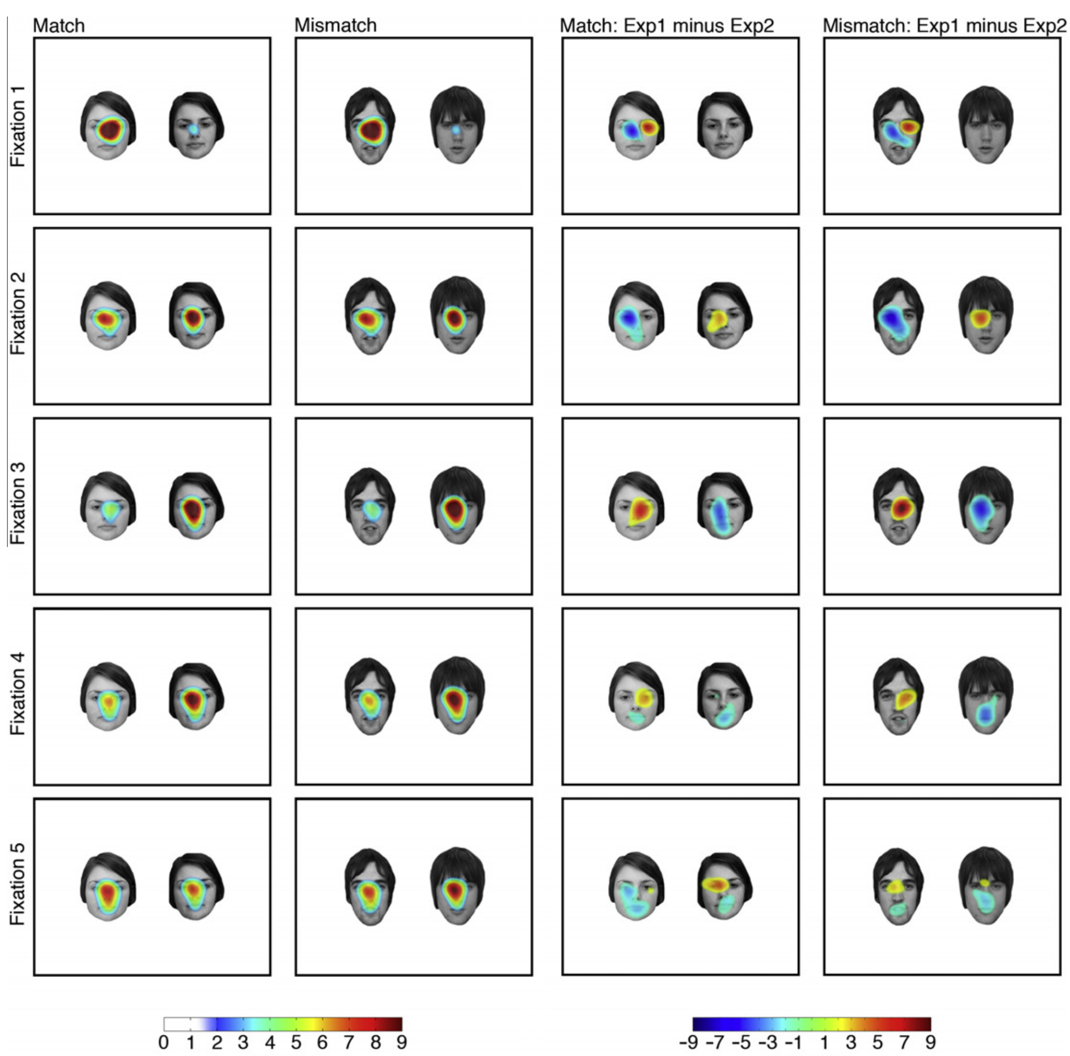

Fig. 8. Gaussian maps illustrating the *z*-scored distribution of fixations in match and mismatch displays in Experiment 2. In addition, difference maps between the distribution of fixations in the unlimited condition of Experiment 1 and Experiment 2 are shown. (For interpretation to colours in this figure, the reader is referred to the web version of this paper.)

short exposure duration and the use of two different photographs of the same person's face for identity *match* trials. The short exposure duration prevents observers from fixating the face pairs and with only a fleeting peripheral view of these stimuli, the identity *matches* might therefore appear to depict two different people. *Match* accuracy then improved rapidly with longer exposure durations, but performance was indistinguishable for 1000 ms, 2000 ms and unlimited displays. This suggests that a combined exposure time for two faces of 1000 ms and a total of 2–3 fixations to these faces is sufficient to elicit optimal accuracy on *match* trials. By comparison, *mismatch* performance was comparable in the 200, 500 and 1000 ms conditions, and peaked later, within 2000 ms. Moreover, except for the 200 ms condition, accuracy was generally lower on *mismatch* than *match* trials. This difference is particularly striking in the unlimited exposure condition as previous studies have shown similar performance for *match* and *mismatch* trials under such unconstrained viewing and with the same face stimuli that were used here (Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010; Megreya, Bindemann, & Havard, 2011).

To explore this discrepancy further, a control study was performed which included only the unlimited display conditions (Experiment 2). This study was essentially a direct replication of previous matching tasks (e.g., Bindemann, Avetisyan, & Blackwell, 2010; Burton, White, & McNeill, 2010), and was conducted to confirm that the effects in Experiment 1 arise from the specific design of the study rather than perhaps some incidental aspects of our general procedure. In this control experiment performance was more accurate on *mismatch* trials in comparison with Experiment 1, and was now also comparable for *match* and *mismatch* displays. Moreover, observers took substantially longer to make these decisions in Experiment 2, which suggests that the face pairs were studied in more detail compared to the unlimited condition in Experiment 1. Taken together, these findings indicate that intermixing different exposure conditions induced a context effect here. Thereby, temporally constrained face viewing also accelerated identification decisions under temporally unconstrained viewing and, as a consequence, reduced *mismatch* accuracy.

These findings are corroborated by the eye movement data, which revealed that different viewing strategies were applied in these experiments. Generally, both experiments yielded a visual field bias, whereby the majority of first fixations were directed to the left face in a pair. This leftward perceptual bias has been observed in studies on many aspects of face perception (see, e.g., Barton et al., 2006; Butler & Harvey, 2008; Coolican et al., 2008; Vinette, Gosselin, & Schyns, 2004), though it appears to reflect a combination of face processing mechanisms (see, e.g., Sergent & Bindra, 1981; Yovel, Tambini, & Brandman, 2008) and scanning habits developed through reading directions (Megreya & Havard, 2011). In Experiment 1, this initial fixation on the left face was followed instantly by a fixation on the face on the right and, if time permitted, by a further switchback to the left. This pattern was found in all conditions (except for 200 ms displays) and, considering that observers had no advance knowledge of how long the faces would be on display, represents an optimal viewing strategy by providing a glance at both faces in a pair with the first two fixations.

In Experiment 1, these fixations landed close to the middle of the face stimuli, but were biased to the side of a face that was closest to the center of the screen display. These landing positions may simply reflect an early scanning effect, whereby observers are initially drawn to the center of mass of a stimulus (e.g., Bindemann, Scheepers, & Burton, 2009; see also Bindemann, 2010; Tatler, 2007). Alternatively, this might represent an optimal viewing position that allows observers to perceive as much of a face as possible (see Hsiao & Cottrell, 2008; Van Belle et al., 2010). Utilizing such an optimal viewing position could facilitate the holistic coding of facial features into an integrated whole, which is held to underlie the identification of familiar faces (e.g., Farah et al., 1998; Maurer, Le Grand, & Mondloch, 2002; Young, Hellawell, & Hay, 1987). Unfamiliar face identification, however, appears to operate best when observers can utilize holistic *and* featural viewing strategies (Van Belle et al., 2010). In line with these findings, observers exhibited a more prolonged encoding strategy in Experiment 2. This viewing strategy also focused initially on the center of the face stimuli but was characterized by slower switching between faces and encompassed more facial features.

Overall, our findings therefore demonstrate that nearly 2 s are required to match two concurrent faces, which equates here to about two fixations *per face*. In this sense, these results converge with claims from recognition memory studies that two fixations may suffice for face identification (Hsiao & Cottrell, 2008). In Hsiao and Cottrell's (2008) study, these measures were obtained after an initial familiarization phase of 3 s and with a subsequent recognition test for the same face *image*. This study therefore provides a very different test for face recognition than the simultaneous matching task of the current experiments, in which identification was assessed across different face photographs (see, e.g., Bruce, 1982; Longmore, Liu, & Young, 2008). These studies converge, however, in suggesting that two glances at a face are sufficient for person identification when viewing conditions are restricted. Moreover, in both of these studies, the faces appeared to be viewed for a similar amount of time when the average fixation duration of $\sim$300 ms in Hsiao and Cottrell's (2008) study is compared, for example, to the 4–5 fixations (including Fixation 0) that observers made within 1.4 s in the 2000 ms condition of Experiment 1 (see Figs. 1 and 2). At the same time, our findings appear to diverge from recognition memory studies that have obtained maximal identification accuracy with extremely short exposures of only 90 ms to a face (Veres-Injac & Schwaninger, 2009). These results were obtained with a small set of faces and a short interval, of only 1000 ms, between test and a more prolonged initial exposure. We therefore attribute recognition with such short exposures to a limited set size, brief retention intervals, and identification across the same face images. While this demonstrates, rather unsurprisingly, that the speed and accuracy of unfamiliar face processing is tied to specific task demands, it also emphasizes the importance of examining the time course of this task under the current conditions.

Intriguingly, the current experiments also suggest that investigations of the time course of unfamiliar face identification are not a straightforward undertaking. This appears to be the case at least with our matching task, in which the intermixing of different exposures produced knock-on effects on the unlimited viewing conditions. As a consequence, accuracy peaked within 1.7 s of stimulus onset even when more time was available to view the faces in the unlimited condition, and this also occurred at a trade-off as *mismatch* accuracy was reduced (cf. Experiments 1 and 2). Intermixing different exposure conditions can therefore provide information about the time course of face matching, but this also appears to distort *match-mismatch* accuracy inadvertently. Moreover, observers' eye movements also revealed that intermixing experimental conditions in this manner affected how faces were viewed, by encouraging observers to switch more quickly between faces and to fixate fewer facial features. Taken together, these findings indicate that exposures of up to 2 s are required to compare pairs of faces when exposure times are variable (Experiment 1), but that such temporal constraints should be avoided altogether if accuracy is truly paramount (Experiment 2).

These findings may be of relevance for some applied settings. Passport control at national borders, for example, typically involves matching a photograph to the face of its purported owner. Psychological research on face matching already suggests that this is an error-prone process under self-paced viewing (e.g., Bindemann & Sandford, 2011; Bindemann, Avetisyan, & Blackwell, 2010; Kemp, Towell, & Pike, 1997). However, in everyday operations that require the routine identification of many people, observers may be unable to study a person's face for as long as is necessary to achieve best-possible accuracy. The current findings suggest that such conditions might lead to an increased acceptance of identity mismatches, even when more time may be available occasionally to identify a face. If this turns out to be the case, then the limited availability of time for *some* person identifications could result in reduced accuracy *all* of the time. This may be an important issue for future investigations.

References

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25*, 997–1010.

Barton, J. J., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception, 35*, 1089–1105.

Bindemann, M. (2010). Scene and screen center bias early eye movements in scene viewing. *Vision Research, 50*, 2577–2587.

Bindemann, M., Avetisyan, M., & Blackwell, K. (2010). Finding needles in haystacks: Identity mismatch frequency and facial identity verification. *Journal of Experimental Psychology: Applied, 16*, 378–386.

Bindemann, M., & Sandford, A. (2011). Me, myself, and I: Different recognition rates for three photo-IDs of the same person. *Perception, 40*, 625–627.

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision, 9*(2), 1–16 (article no. 17).

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE, 3*(8), e3022.

Bruce, V. (1982). Changing faces: Visual and non-visual coding processes in face recognition. *British Journal of Psychology, 73*, 105–116.

Burton, A. M., White, D., & McNeill, A. (2010). The Glasgow face matching test. *Behavior Research Methods, 42*, 286–291.

Butler, S. H., & Harvey, M. (2008). Effects of aging and exposure duration on perceptual biases in chimeric face processing. *Cortex, 6*, 665–672.

Coolican, J., Eskes, G. A., McMullen, P. A., & Lecky, E. (2008). Perceptual biases in processing facial identity and emotion. *Brain and Cognition, 66*, 176–187.

Davies, G., & Griffiths, L. (2008). Eyewitness identification and the English courts: A century of trial and error. *Psychiatry, Psychology, and Law, 15*, 435–449.

Duchaine, B. C., & Nakayama, K. (2004). Developmental prosopagnosia and the Benton facial recognition test. *Neurology, 62*, 1219–1220.

Duchaine, B. C., & Weidenfeld, A. (2003). An evaluation of two commonly used tests of unfamiliar face recognition. *Neuropsychologia, 41*, 713–720.

Ellis, H. D., Young, A. W., & Koenken, G. (1993). Covert face recognition without prosopagnosia. *Behavioural Neurology, 6*, 27–32.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review, 105*, 482–498.

Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science, 198*, 853–855.

Hallett, P. E. (1986). Eye movements. In K. B. Boff, L. Kaufman, & J. P. Thomas (Eds.). *Handbook of perception and human performance: Sensory processes and perception* (Vol. I, pp. 10–102). New York: Wiley.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*, 498–504.

Henderson, J. M. (2007). Regarding scenes. *Current Directions in Psychological Science, 16*, 219–222.

Henderson, J. M., Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory and Cognition, 33*, 98–106.

Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science, 19*, 998–1006.

Janik, S. W., Wellens, A. R., Goldberg, M. L., & Dell'Osso, L. F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual and Motor Skills, 47*, 857–858.

Johnston, R. A., Hill, H., & Carman, N. (1992). Recognizing faces: Effects of lighting direction, inversion, and brightness reversal. *Perception, 21*, 365–375.

Kemp, R., Towell, N., & Pike, G. (1997). When seeing should not be believing: Photographs, credit cards and fraud. *Applied Cognitive Psychology, 11*, 211–222.

Logie, R. H., Baddeley, A. D., & Woodhead, M. M. (1987). Face recognition, pose and ecological validity. *Applied Cognitive Psychology, 1*, 53–69.

Longmore, C. A., Liu, C. H., & Young, A. W. (2008). Learning faces from photographs. *Journal of Experimental Psychology: Human Perception and Performance, 34*, 77–100.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Science, 6*, 255–260.

Megreya, A. M., Bindemann, M., & Havard, C. (2011). Sex differences in unfamiliar face identification: Evidence from matching tasks. *Acta Psychologica, 137*, 83–89.

Megreya, A. M., & Havard, C. (2011). Left face matching bias: Right hemisphere dominance or scanning habits. *Laterality, 26*, 75–92.

Meinhardt-Injac, B., Persike, M., & Meinhardt, G. (2010). The time course of face matching by internal and external features: Effects of context and inversion. *Vision Research, 50*, 1598–1611.

Morrison, D. J., Bruce, V., & Burton, A. M. (2000). Covert face recognition in neurologically intact patients. *Psychological Research, 63*, 83–94.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice-Hall.

Scheck, B., Neufeld, P., & Dwyer, J. (2000). *Actual innocence*. New York: Random House.

Sergent, J., & Bindra, D. (1981). Differential hemispheric processing of faces: Methodological considerations and reinterpretation. *Psychological Bulletin, 89*, 541–554.

Simons, D. J., & Levin, D. T. (1998). Failure to detect changes to people in real-world interaction. *Psychonomic Bulletin and Review, 5*, 644–649.

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision, 7*(14), 1–17 (article no. 4).

Van Belle, G., De Graef, P., Verfaillie, K., Busigny, T., & Rossion, B. (2010). Whole not hole: Expert face recognition requires holistic face perception. *Neuropsychologia, 48*, 2620–2629.

Veres-Injac, B., & Schwaninger, A. (2009). The time course of processing external and internal features of unfamiliar faces. *Psychological Research, 73*, 43–53.

Vinette, C., Gosselin, F., & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: It's in the eyes. *Cognitive Science, 28*, 289–301.

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological Science in the Public Interest, 7*, 45–75.

Young, A. W., Hellawell, D., & de Haan, E. H. F. (1988). Cross-domain semantic priming in normal subjects and a prosopagnosic patient. *The Quarterly Journal of Experimental Psychology, 40A*, 561–580.

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception, 16*, 747–759.

Yovel, G., Tambini, A., & Brandman, T. (2008). The asymmetry of the fusiform face area is a stable individual characteristic that underlies the left-visual-field superiority for faces. *Neuropsychologia, 46*, 3061–3068.